开放科学创新生态构建中的数据处理伦理准则框架研究

刘静羽[1,2]，贾毓洁[1,2*]，黄金霞[1]，王　昉[1]

（1.中国科学院文献情报中心，北京　100190；2. 中国科学院大学 经济与管理学院图书情报与档案管理系，北京　100190）

摘　要：[目的／意义]开放科学发展仍面临着诸多数据伦理难题，亟需探究与开放科学创新生态相适应的数据处理伦理准则。建立与开放科学创新生态构建相匹配的数据处理伦理准则，指导开放科学创新生态中的参与主体在数据创造、共享、传播、利用过程中按照法律规定和伦理规范开展行动，约束数据处理行为向好向善、造福社会，对于营造良好开放科学创新生态至关重要。[方法／过程]本研究首先运用文献调研、内容分析法，立足数据处理伦理的相关法律、指南、框架草案、政策以及典型数据管理模型，从普适的角度出发，立足数据获取、存储、管理、使用和销毁的实践，面向广义的社会性数据，同时兼顾一部分科研数据，关联整合不同的准则内容，寻求现有准则间的共有主题和内容，归纳数据处理伦理准则的一般维度和认可度较高的准则内容。其次，本研究结合开放科学理念和国家社会发展需要，从国内最新规划政策与开放科学内涵中，提炼出开放科学创新生态环境下的数据处理伦理准则需求。最后，本研究通过矩阵分析综合比较两项结果，从共性和特性、准则内容与政策及理念协同等角度，分析得出开放科学创新生态构建中的数据处理伦理准则框架。[结果／结论]现有的数据处理伦理准则在开放科学创新生态中具有较高的适用性，但仍需动态拓展。本研究结合开放科学理念和国家社会发展需要，提出开放科学创新生态数据处理伦理准则框架，包括长治久安、协调有序和系统科学 3 个维度和 15 个要素，旨在服务于适应中国开放科学创新生态发展的数据处理伦理准则文本的进一步制定，并为机构的数据管理提供一定参考。

关键词：数据处理伦理；开放科学创新生态；开放科学；数据治理；科技伦理

中图分类号：G250.7　　　**文献标识码：**A　　　**文章编号：**1002-1248（2022）09-0029-15

引用本文：刘静羽，贾毓洁，黄金霞，等. 开放科学创新生态构建中的数据处理伦理准则框架研究[J]. 农业图书情报学报，2022，34(9)：29-43.

1 引　言

开放科学是全球科技发展的重要趋势，构建“开放、信任、合作”的开放科学创新生态被认为是推动开放科学的重要途径[1]。当前，数据作为社会新型生产要素[2]和战略性资源[3]，与开放科学的演进深度交融，以数据复用驱动的科研基础设施建设成为中国现阶段

收稿日期：2022-06-22

基金项目：中国科学院文献情报能力建设专项“建设全球高质量开放学术论文汇聚平台”（E1291902）；国家科技图书文献中心专项任务“开放资源权益及版权谈判与规范应用维护”（2022XM12）

作者简介：刘静羽（1989- ），硕士，副研究馆员，研究方向为开放资源建设、预印本学术交流、数字资源权益管理。黄金霞（1972- ），博士，研究馆员，研究方向为开放资源建设、开放科学研究。王昉（1977- ），硕士，副研究馆员，研究方向为开放资源建设与利用。

* 通信作者：贾毓洁（1997- ），硕士研究生，研究方向为数字图书馆技术、数据管理、数据权益。Email：jiayujie@mail.las.ac.cn

开放科学实践的核心[4]。习近平总书记在 2021 年 6 月 30 日的两院院士大会中国科协十大重要讲话中提出要"构建开放创新生态，参与全球科技治理"[5]，突出强调了前瞻预判科技发展中的伦理挑战以及要进一步完善伦理审查规则及监管框架。

做好数据治理，建立与开放科学创新生态构建相匹配的数据处理伦理准则，指导开放科学创新生态中的参与主体在数据创造、共享、传播、利用过程中按照法律规定和伦理规范开展行动，约束数据处理行为向好向善、造福社会，对于营造良好开放科学创新生态至关重要。

2 数据处理伦理与开放科学

2.1 数据处理伦理

伦理，是建立在是非观念上的行为准则，伦理准则侧重于公平、尊重、责任、诚信、质量、可靠性、透明度和信任等方面，主要包括道德准则、法律规范和利益分配等内容。2020 年 9 月，美国总务署（General Services Administration，GSA）发布的《数据伦理框架草案》中，数据伦理概念被定义为在收集、管理或使用数据时，为保护公民自由、最大限度地降低个人和社会的数据使用风险、实现公共利益最大化等目的，进行适当判断和问责的依据。数据处理伦理，是指如何以符合伦理准则的方式获取、存储、管理、使用和销毁数据，是从实践出发的狭义数据伦理概念，主要应用于组织机构数据管理层面[6]。

数据生命周期的各个阶段，包括数据产生、获取、存储、传播、利用等活动，都需要数据处伦理的指导，以保证数据质量。明确数据处理伦理准则，能有效地规范组织机构对数据的管理与利用。基于伦理准则处理数据，对于任何希望从数据中持续获得价值的组织和个体来说都是必要的；反之，违反数据处理伦理准则会导致组织面临声誉损失、用户流失等风险，一些违反伦理的行为甚至会触犯法律。

随着社会探索和利用大量数据和信息的专业化能力不断提升，数据伦理与人类社会生产生活紧密融合，重要性日益增强。数据处理伦理已被纳入国际公认和权威的数据管理知识体系中，在 2017 年最新出版的数据管理领域权威性基础工具书《DAMA 数据管理知识体系指南》中被作为单独的一章加以论述。国际标准化组织（ISO）、国际数据治理研究所（DGI）、Gartner 公司、国际信息系统审计和控制协会（ISACA）等均尝试构建数据治理模型 / 框架，并强调以伦理规范指导数据管理，控制数据质量，防范数据被歪曲、滥用和误解的风险。

2.2 开放科学环境下的数据伦理难题

在数据管理领域，数据处理伦理在理论和实践层面都已经有了一定程度的发展和应用。但随着开放科学在中国的持续推进，开放科学创新生态构建成为未来发展的大趋势，其对数据治理的要求与数据处理伦理准则的核心要义不谋而合。

但数据处理伦理是一个较为复杂的问题，不仅横跨生物医学、空间科学、教育管理、计算机科学、新闻传播等多个学科，还关系到政府、企业、个人等利益相关主体。中国数据要素市场尚不完善，智能技术应用[7]中存在数据权属不明确、数据标准不一、数据共享开放不充分[8]等伦理和道德问题，阻碍了数字化时代中国的开放科学发展。刘金亚等[9]分析并得出了 23 个开放科研数据环境下的伦理问题，覆盖数据创建者对数据的控制权、数据描述规则、数据滥用、数据透明、知情同意、数据隐私等多个方面；陈劲等[10]提出数据真实性审查，以及基于机器学习、大数据分析的智能数据采集、智能决策可能带来的负面社会影响均属于开放科学须直面的伦理问题；李伦[11]指出人与自由的关系是数据伦理处理准则的核心，数据伦理问题早已随着大数据技术的发展而受到广泛关注，成为技术伦理学探寻的热点命题之一。

因此，有必要研究和探讨开放科学创新生态构建中的数据处理伦理问题，包括：在开放科学创新生态构建中各利益主体需要遵循的数据处理伦理准则是什么；各国和各组织机构间已有的数据处理伦理准则是

否可以移植到开放科学创新生态环境；开放科学创新生态中的数据处理伦理准则是否对原有准则内容有新的内涵或扩展。

本研究运用文献调研、内容分析法，从相关法律、指南、框架和典型数据管理模型中，调研归纳数据处理伦理准则的一般维度和认可度较高的准则内容，并从国内最新规划政策与开放科学内涵中，提炼出开放科学创新生态环境下的数据处理伦理准则需求。之后，通过矩阵分析综合比较两项结果，从共性和特性、准则内容与政策及理念协同等角度，分析得出开放科学创新生态构建中的数据处理伦理准则框架。该框架未来可用于制定适应中国开放科学创新生态构建的数据处理伦理准则文本，指导各相关利益主体形成对伦理的准确认识，使其在开展数据处理活动时，规范数据处理方式和工作流程，甄别和规避伦理风险，进而推动中国开放科学创新生态的可持续发展。

3　数据处理伦理准则的内容及维度分析

数据处理伦理仍在发展中，尚未形成一套统一的框架。受到法律、组织文化差异等影响，现有的数据处理伦理准则在国家和组织机构之间呈现出一定的内容差异性和分布零散性，不同的准则有各自的侧重。本研究从普适的角度出发，立足数据获取、存储、管理、使用和销毁的实践，面向广义的社会性数据，同时兼顾一部分科研数据，关联整合不同的准则内容，寻求现有准则间的共有主题和内容[12]。

3.1　数据处理伦理文本选取

伦理准则从类型来讲，覆盖“硬法”（强制性的法律和条约）、“软法”（志愿性的、非约束力的协议或指南）和非正式措施（行为准则和提高意识等道义劝告）；从作用面来讲，又可分为国际、区域、国家、行业等多个尺度[13]。从数据处理伦理的普适性出发，本研究选取目前国际上普遍认可、实践中较为领先的数据处理伦理相关法律法规、报告指南、框架草案及政策，同时也关注了模型中的数据处理伦理准则，期望从中提取和梳理数据处理伦理准则的基本维度。

3.1.1　法律层面的强伦理约束

随着数据在社会生产生活中的愈来愈占据重要地位，各国陆续制定了一系列法律法规，对数据处理活动进行规制。这些固定性、强制性的规则部分涉及由“隐私”“安全”等道德伦理上升而来的内容，为数据处理相关的活动划下了一道红线，是社会生活中的数据处理活动最基本的强伦理约束。

欧盟 2018 年发布的 GDPR，一以贯之个人隐私保护的原则，是目前全球在保护个人数据方面，规定最为严格、处罚最为严厉的法规；加拿大响应个人信息保护国际运动号召，2019 年发布了 PIPEDA，为通过商业渠道采集的欧盟国家公民的个人信息提供了强有力的保护；美国作为世界上最早提出并通过法规对隐私权予以保护的国家，1974 年的《隐私法案》是最重要的一部保护个人信息方面的法律，2019 年的 NSPDPA 则开始限制跨境数据流向，将数据保护提升到国家安全层面；中国则在“十四五”开局之际，相继完善和推出了《数据安全法》《个人信息保护法》《数据出境安全评估办法》《网络数据安全管理条例》等多项法律法规，进入数据安全保护的法制化时代。

这些数据处理相关的代表性法律法规（表 1），可以在一定程度上反映出不同区域或国家对于数据处理所秉持的态度、最基本的要求和底层的道德伦理逻辑。

3.1.2　指导层面的弱伦理约束

尽管部分伦理道德可以上升为法律，但法律无法完全代替道德伦理，法律的固定性和强制性决定了其在快速发展变化的社会环境下，一定程度上滞后于现实需要。因此，面对层出不穷的数据处理伦理需求，各行各业陆续制定公布了数据处理相关的报告指南、框架草案以及政策，从而进行自我规制。

宏观上，1979 年的《贝尔蒙报告》提出的贝尔蒙特医学研究原则也适用于信息管理学科，2012 年的《Menlo 报告》则将贝尔蒙特医学准则用于信息和通讯技术研究，为数据伦理准则提供了一个良好的起点；《大数据的数据隐私、道德和保护指导说明》和《DATAETHICS—公司、当局和组织的原则和指南》与

表 1 数据处理伦理相关的法律法规

Table 1 Laws and regulations related to data handling ethics

编码	题名	介绍	责任方
A1	通用数据保护条（GDPR）[14]	约束欧盟境内的机构组织对个人信息的收集、传输、保留或处理	欧盟
A2	个人信息保护和电子文件法案（PIPEDA）[15]	约束加拿大各省涉及个人信息收集、使用和披露的商业活动，以及个人信息在国际和省之间的传递	加拿大
A3	美国隐私法案（1974）[16]	规范美国政府机构收集、维护、使用和传播个人信息的最重要的一部法律，规定了信息主体的权利	美国
A4	国家安全与个人数据保护法案（NSPDPA）[17]	以保护本土居民企业和国民数据为切入口，限制跨境数据流向	美国
A5	中华人民共和国数据安全法[18]	数据安全管理的基本大法，指导有关单位和个人收集、存储、使用、加工、传输、提供、公开数据资源	中国
A6	中华人民共和国个人信息保护法[19]	加强个人信息保护法制保障、维护网络空间良好生态、促进数字经济健康发展	中国
A7	数据出境安全评估办法[20]	规定了数据出境时需要进行安全评估的情形、具体内容、提交材料、评估事项、评估流程等	中国
A8	网络数据安全管理条例[21]	规定了数据处理者使用及获取数据、数据处理者本身发生主体变化、数据跨境等情形时应当遵循的数据处理规则	中国

数据隐私、数据保护和数据道德相关，可以帮助指导将数据伦理整合到数据处理活动中。微观上，《EDPS道德框架》《数据伦理框架》《伦理和数据保护》《数据伦理框架草案》《数据处理伦理语境关系》以及《个性化健康研究中负责任数据处理的伦理框架》则从具体实践出发，为组织和利益相关方的数据处理和操作提供了伦理参考依据。更有《科学数据管理办法》《开放存取与数据传播和保存政策指南》《开放数据白皮书：释放潜能》《开放数据政策——将信息作为资产进行管理》《NIH 数据管理和共享最终政策》为国家、机构、组织等的数据管理和开放共享提供方向指导。

这些数据处理相关的代表性报告指南、框架草案及政策（表 2），是各领域和行业从业人员对于其数据处理伦理需求的补充，更具体地反映出从业人员对于数据处理行为规范有着更高级别的要求。

3.1.3 实践层面持续发挥作用的伦理约束

数据处理伦理还在数据管理过程中持续发挥着指导作用。数据管理是一个复杂的过程，是一项跨职能的工作[6]，数据质量保证这一难题贯穿着数据管理的整个生命周期[36]。在数十年的时间，DataFlux、Gartner、IBM、斯坦福大学、企业数据管理委员会（EDM Council）、卡耐基 - 梅隆大学旗下的 CMMI 协会以及中国国家标准管理委员会等国际组织、跨国 IT 服务企业和权威咨询机构相继推出模型以从不同视角揭示数据管理过程中的重要内容（表 3），其中数据质量管控、数据安全策略、信息隐私保障、数据文化构建、信息风险防范等目标都与数据处理伦理相关。

3.2 数据处理伦理维度层级

从上述数据来源和内容可以看出：数据处表 3 理伦理，以法律法规为准绳向上下延伸，向上是宏观的道德伦理准则，向下是微观的数据操作细则要求（图 1）。

数据处理伦理相关的法律法规以保障国家安全、维护社会和公众的利益为核心原则，为数据处理相关的活动划下了一道红线，为组织保护数据提供了动机。但一些组织认为只要遵守法律相关规定，就不会有数据处理相关风险，这是一个危险的假设，因为立法往往跟不上数据生态变化带来的风险，故而在法律法规的指导下，还需要从宏观角度理解数据处理的道德伦理准则，从微观视角规范数据处理操作细节要求。从宏观来看，数据处理伦理的原则包括维护公共安全、保护社会公平、尊重主体权利、重视隐私保护、推动

表 2　数据处理伦理相关的报告指南、框架草案及政策

Table 2 Reporting guidelines, draft frameworks and policies related to data handling ethics

类型	编码	题名	介绍	责任方
报告指南	B1	贝尔蒙报告（US-HSS）[22]	贝尔蒙特医学研究原则也适用于信息管理学科，为数据伦理准则提供了一个良好的起点	美国
	B2	Menlo 报告（US-DHS）[23]	将贝尔蒙特医学准则用于信息和通讯技术研究	美国
	B3	大数据的数据隐私、道德和保护指导说明[24]	制定了关于使用大数据的数据隐私、数据保护和数据道德的一般指南	联合国
	B4	DATAETHICS——公司、当局和组织的原则和指南[25]	制定了一套数据伦理原则和指南，可以帮助将数据伦理整合到数据处理活动中	DataEthics（丹麦）
框架草案	C1	EDPS 道德框架[26]	从所有员工的行为规范出发，呼吁关注维护人类尊严的数据处理	欧盟
	C2	数据伦理框架[27]	指导公共部门组织在规划、实施和评估新政策或服务时适当和负责任地使用数据	英国
	C3	伦理和数据保护[28]	指导以符合法律和道德要求的方式进行数据收集和处理操作	欧盟
	C4	数据伦理框架草案[29]	帮助机构雇员、经理人和领导人在获取、使用数据时合乎伦理要求	美国
	C5	数据处理伦理语境关系[6]	描述了数据处理伦理的领域知识结构	DAMA 国际
	C6	个性化健康研究中负责任数据处理的伦理框架[30]	确保在 SPHN 范围内进行的与个人数据和人类生物材料相关的科学活动符合适当的道德责任标准	瑞士
政策	D1	科学数据管理办法[31]	要求科学数据管理遵循分级管理、安全可控、充分利用的原则，明确责任主体，加强能力建设，促进开放共享	中国
	D2	开放存取与数据传播和保存政策指南[32]	分析了开放存取和数据传播和保护相关的广泛法律和道德挑战，包括知识产权和数据保护等	欧盟
	D3	开放数据白皮书：释放潜能[33]	要求政府部门以机器可读格式发布政府数据，建立透明和开放的政府	英国
	D4	开放数据政策——将信息作为资产进行管理[34]	要求各机构以支持下游信息处理和传播活动的方式收集或创建信息，包括使用机器可读和开放格式、数据标准、可扩展元数据等，最大限度地提高数据的互操作性和信息可访问性	美国
	D5	NIH 数据管理和共享最终政策[35]	确立了最大限度公开和共享由 NIH 资助或开展的科研项目所产生的科研数据	美国

开放共享、保障数据生态的可持续发展；从微观来看，在数据处理的各个过程中，相关的操作要求要遵循法律规范和伦理原则进一步细化，例如：数据收集的知情同意、数据存储的安全保障、敏感数据的脱敏处理、责任主体的合理界定、数据管理的信任机制等。

3.3　数据处理伦理内容分析

以上述选定的文本为研究对象，结合数据处理伦理的维度层级，本研究将对数据处理伦理进行内容分析，主要包括 3 个步骤：概念化、范畴化和逻辑整合。概念化是对原始文本语句依据普遍理解，通过贴标签后形成概念；范畴化是对已形成概念的归纳整合；逻辑整合则是在已有概念和范畴的基础上深挖其中的逻辑关系，形成有意义的理论。

3.3.1　概念化和范畴化

本研究针对原始文本，并未严格按照逐词、逐句、

表 3 数据处理伦理相关的数据管理模型

Table 3 Data management models related to data handling ethics

编码	模型	数据处理伦理相关的描述	责任方
M1	DataFlux Maturity Model[37]	防范数据问题带来的成本和商誉损失；重视数据质量控制，通过高质量的数据驱动企业决策	DataFlux
M2	EIM Maturity Model[38]	与信息有关的责任是在逐个应用程序和逐个项目的基础上分配的；数据质量已成为文化的一部分；确保信息安全，防范信息风险	Gartner
M3	数据治理能力成熟度模型[39]	以增强企业数据的可用、可信、可管、可控为目标；以数据质量管理、信息生命周期管理、信息安全与隐私为核心准则	IBM
M4	数据治理成熟度模型[40]	建立数据文化；培养数据思维；实施数据质量控制	斯坦福大学数据治理中心
M5	CMMI-DMM[41]	从质量战略、质量标准、质量规则等方面支撑数据质量	CMMI
M6	DCAM[42]	通过数据反映业务流程和操作的程度，如准确性、完整性、及时性等，贯穿数据生命周期全过程，确保数据满足预期利益相关者需求。	EDM Council
M7	DCMM[43]	数据安全主要包括数据安全策略、数据安全管理和数据安全审计等；数据质量主要是指数据的准确性、及时性、完整性、唯一性、一致性、有效性等	中国标准化管理委员会

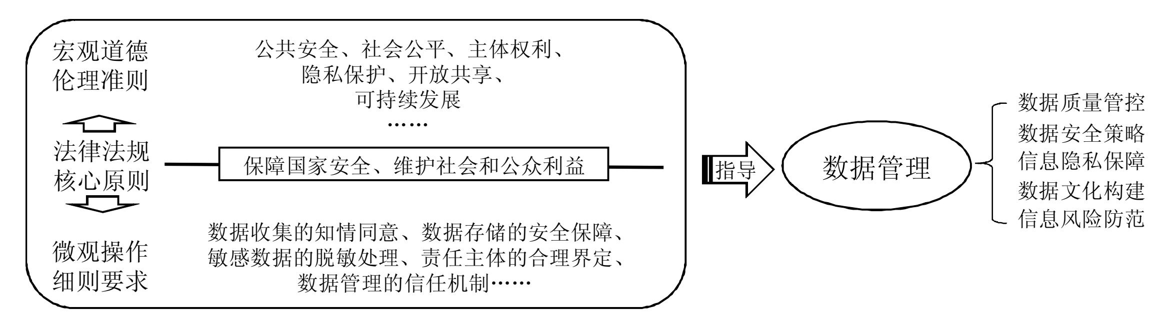


图 1 数据处理伦理维度层级结构

Fig.1 Dimension hierarchy of data handling ethics

逐行或逐段的某种固定方式提取概念，而是选取可操作的、有意义的某些片段进行概念化和范畴化，选取的内容主要包括两个方面：一是隐性体现的一般性伦理准则；二是显性提出的一般性伦理准则。由于后者一般是已经形成的概念，故而选取的资料还是以前者为主，后者多作为辅助材料在概念化的过程中作为参考，部分概念化过程如表 4 所示。

3.3.2 逻辑整合

通过对选定的文本进行概念化和范畴话后，经过归纳整合，最后共得到数据处理伦理相关的 11 个关键要素，包括：公众安全、个人主权、隐私保护、数据安全性、数据准确性、数据时效性、数据可信度、数据透明度、人员培训、监管机制和风险防控。将这些要素依据宏观、中观和微观 3 个层面进行归纳整合后，可以得到数据处理伦理的 3 个主题维度：长治久安、协调有序以及系统科学（图 2）。

（1）以长治久安为目标，体现国家安全观和社会责任观。从国家安全观和社会责任观看，数据处理伦理以长治久安为目标，包括公众安全、个人主权和隐私保护。公众安全包括公众利益和社会安全，是数据处理伦理相关法律法规的核心原则，也是数据利用的终极目标。个人主权即尊重人的权利。大数据时代，代表个人特征的数据一旦被用于决策时，会影响人们的生活，故而无论是数据生产者还是数据利用者，都要以“不伤害”的原则加强对个人权利的尊重和保护。隐私保护是数据处理伦理中被重点强调的内容，在收集、处理和分析他人产生的数据时要注重隐私保护，这既是对数据生产者的尊重，也是数据处理者保护数据安全的重要一环。

（2）以协调有序为方向，强调规范秩序和行业自

表 4　概念化和范畴化过程示例（部分）

Table 4　Example of conceptualization and categoryization process (partial)

范畴化	概念化	文本材料
公众安全	[1-1]尊重法律和公共利益（a1,…）	a1：尊重法律和公共利益[B2]
	[1-2]维护国家主权、安全（a2,a3,…）	a2：维护国家主权、安全和发展利益[A5]
	[1-3]保障公共利益（a5,a8,c4,…）	a3：维护国家安全、公共利益[A8]
	[1-4]产生公众信任（c6,…）	a4：保障公共利益[C4]
	……	a5：确保有效生产有价值的科学知识，并围绕网络活动产生公众信任[C6]
		……
个人主权	[2-1]尊重个人权益（b1,b2,b3,…）	b1：尊重他人、行善原则、公正[B1]
	[2-2]以人为本（b1,c3,…）	b2：保护个人、组织的合法权益[A5]
	[2-3]合法正当（b2,b4,b6,…）	b3：促进研究参与者的权利、利益和福祉[C6]
	[2-4]公平公正（b1,…）	b4：保护个人、组织在网络空间的合法权益[A8]
	……	b5：任何组织、个人不得侵害自然人的个人信息权益[A6]
		b6：处理个人信息应当遵循合法、正当、必要和诚信原则[A6]
		b7：数据保护既是欧洲研究伦理的核心问题，也是一项基本人权[C3]
		……
……	……	……
风险防控	[11-1]风险模型（k1,…）	k1：伦理风险模型[C5]
	[11-2]风险管理（k2,k3,…）	k2：数据风险管理，数据风险管理涉及的风险识别、量化、规避、接受和减轻等[M3]
	[11-3]风险防范（k4,…）	k3：数据管理，旨在确保组织获得高质量数据资产，开展行之有效的数据管理，以降低风险和提升数据资产利用价值[M3]
	……	k4：防范数据出境安全风险，保障数据依法有序自由流动[A7]
		风险缓解和风险、危害和收益评估[B3]
		……

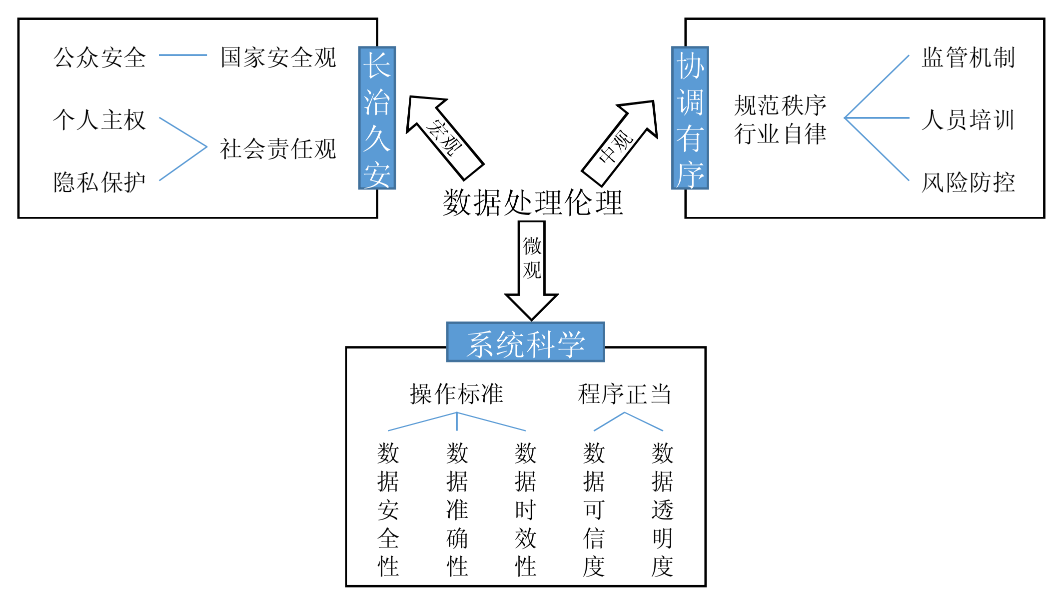


图 2　数据处理伦理维度整合

Fig.2　Dimension integration of data handling ethics

律。在规范秩序和行业自律方面，数据处理伦理以协调有序为方向，包括人员培训、监管机制和风险防范。对于组织机构来说，对人员进行培训，加强伦理观念和伦理准则，是提高数据素养的直接途径。同时，建立数据处理和监督的方法、风险防范相关制度，定期进行数据审计，也是确保数据被正确使用的重要措施，例如 DAMA 国际建立伦理风险模型工具，有助于帮助组织及时发现问题并解决。

（3）以系统科学为指导思想，要求操作标准和程序正当。在具体实践方面，数据处理伦理以系统科学为指导思想，要求操作标准和程序正当，包括数据安全性、数据准确性、数据时效性、数据可信度和数据透明度。这些伦理准则体现为一系列规范措施，例如数据保障措施、目的限制、问责制度、授权许可以及合规挑战等方面的规章制度，其目标是确保数据不被滥用以及不被用于欺骗和误导，以保证数据分析和数据支撑决策的精准性，从而规避数据风险导致的成本和信誉损失，最终帮助组织积极承担责任，构建良好的数据生态。

4 开放科学创新生态中的数据处理伦理准则框架

开放科学创新生态，是开放科学发展与开放创新生态建设的融合，体现着科学与社会的衔接。本研究在前文所描述的数据处理伦理准则维度基础上，结合开放科学理念和国家社会发展需要，从国内最新规划政策与开放科学内涵中，提炼出开放科学创新生态环境下的数据处理伦理准则需求后，通过矩阵分析综合比较两项结果，从共性和特性、准则内容与政策及理念协同等角度，分析得出开放科学创新生态构建中的数据处理伦理准则框架。

4.1 开放科学创新生态发展总体目标与基本原则

开放科学创新生态的构建不能脱离国家和社会发展，开放科学创新生态中的数据处理伦理准则也必须与国家科技创新、数字化转型发展中的理念、原则相契合。本研究选取《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》，以及 2021 年出台或实施的部分法律和规划，分析其总体目标、基本原则、与数据处理相关内容等，提取相应关键词，为设计开放科学创新生态中数据处理伦理准则的参考和依据（表 5）。

上述所有理念的并集，应成为开放科学创新生态数据处理伦理准则的指导与依据，而表中显示的“安全、质量、开放、诚信、系统”等高频关键词，应在开放科学创新生态数据处理伦理准则设计中得到更为突出的体现。例如，在数据处理全过程中，强化数据安全保障责任和风险防控意识，防范数据泄露，注重关乎国家科技安全、企业发展的重要数据保护，以及个人信息和隐私数据保护，重视数据记录、溯源等；尊重数据所有权、使用权、控制权、交易权等多种权益，并积极根据国家法律法规、数据要素市场发展、数据产权框架发展的动态变化，调整数据处理伦理准则中关于数据权属的规定；强调数据处理主体间的协同高效、开放合作等理念，提高行业的自律意识，营造良好的数据伦理生态。

4.2 开放科学创新生态中的数据处理伦理准则框架构建

由上述分析可知，开放科学理念与数据处理伦理准则一般维度相契合。故而本研究在数据处伦理一般维度的基础上，根据开放科学理念和国家社会发展需要所提炼和增加的内容，如图 3 中蓝色部分的要素，构建起开放科学创新生态中的数据处理伦理准则框架，共包含 3 个维度、15 个要素。

4.3 开放科学创新生态中的数据处理伦理准则框架发展与应用

根据联合国教科文组织发布的《开放科学倡议书》，开放科学的理念是“参与、包容、分享、合作、公开、透明”[44]，这与数据处理伦理准则中“长治久安、协调有序、系统科学”3 个维度是高度契合的，开

表 5 国内部分法律政策分析及关键词提取

Table 5 Analysis and keyword extraction of some domestic legal policies

名称	总体目标/基本原则	关键词
中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要	基本实现国家治理体系和治理能力现代化，人民平等参与、平等发展权利得到充分保障，基本建成法治国家、法治政府、法治社会	平等、法治
中华人民共和国科学技术进步法	国家统筹发展和安全，提高科技安全治理能力，健全预防和化解科技安全风险的制度机制，加强科学技术研究、开发与应用活动的安全管理，支持国家安全领域科技创新，增强科技创新支撑国家安全的能力和水平	安全、风险
中华人民共和国数据安全法	开展数据处理活动，应当遵守法律、法规，尊重社会公德和伦理，遵守商业道德和职业道德，诚实守信，履行数据安全保护义务，承担社会责任，不得危害国家安全、公共利益，不得损害个人、组织的合法权益	法律、公德、道德、诚信、安全、责任
中华人民共和国个人信息保护法	处理个人信息应当遵循合法、正当、必要和诚信原则，不得通过误导、欺诈、胁迫等方式处理个人信息。 处理个人信息应当遵循公开、透明原则，公开个人信息处理规则，明示处理的目的、方式和范围。处理个人信息应当保证个人信息的质量，避免因个人信息不准确、不完整对个人权益造成不利影响	合法、正当、必要、诚信、公开、透明、质量
“十四五”大数据产业发展规划	价值引领,基础先行,系统推进,融合创新,安全发展,开放合作	价值、基础、系统、融合、创新、安全、开放
“十四五”数字经济发展规划	创新引领、融合发展，应用牵引、数据赋能，公平竞争、安全有序，系统推进、协同高效	创新、整合、应用、赋能、公平、安全、有序、系统、协同、高效
“十四五”信息化和工业化深度融合发展规划	市场主导、创新驱动、系统推进、开放合作	市场、创新、系统、开放、合作
“十四五”国家信息化规划	坚持党的领导、坚持以人民为中心、坚持新发展理念、坚持深化改革开放、坚持系统推进、坚持安全与发展并重	安全、发展、系统

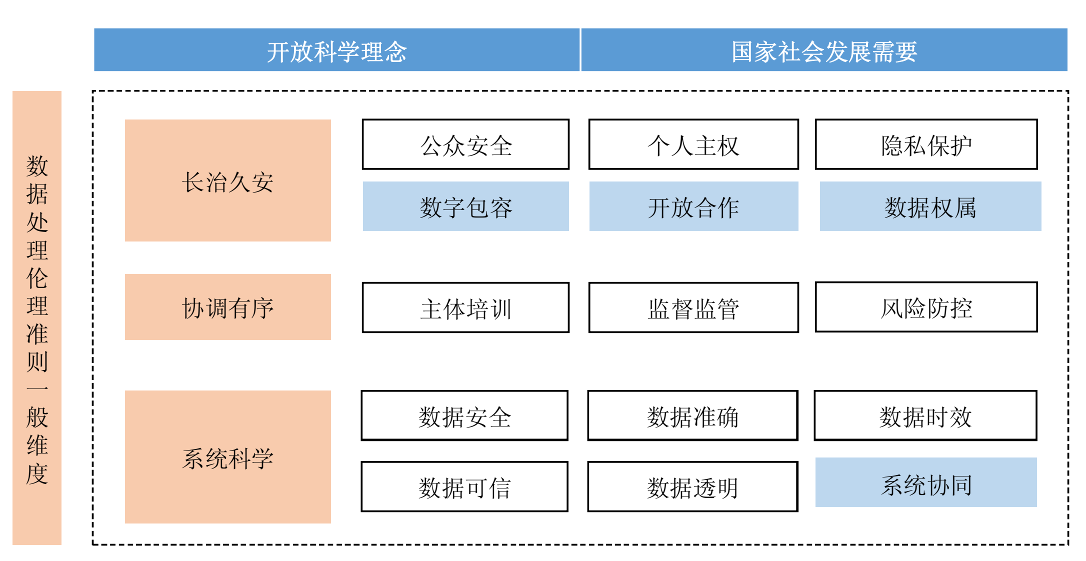


图 3 开放科学创新生态中的数据处理伦理准则框架

Fig.3 Framework for data handling ethical guidelines in the open science innovation ecology

放科学创新生态建设应从以下5个方面关注数据处理伦理。

（1）科学研究是造福人类的事业，须向善而行。正如2017年欧盟特别峰会所签署的《罗马宣言》指出的“科研和创新的决策尊重人的尊严、自由、民主、平等、法治及尊重人权”。2021年12月中央全面深化改革委员会第二十三次会议审议通过《关于加强科技伦理治理的指导意见》也强调，科技活动必须遵守的科技伦理准则，包括坚持增进人类福社、尊重生命权利、公平公正、合理控制风险、公开透明等[45]。因此，开放科学中的数据处理必须以尊重人权为首要原则，同时重视保护隐私权。例如在有关临床病例或者环境监测的公民科学研究中，往往涉及个人数据的收集和身份信息的识别，需要加强对隐私的保护。

（2）开放科学中的数据处理伦理准则应更多体现包容性。联合国2030可持续发展议程提出17个可持续发展目标，其中有5个都提及“包容”。联合国教科文组织也在积极探讨如何通过开放科学推动实现更大范围的数字包容，弥补仍然存在且可能进一步加剧的数字鸿沟。目前主要发达国家的数据治理中都已考虑数字包容问题[46]，中国在构建开放科学创新生态过程中，也应考虑将“包容”原则纳入数据处理的伦理准则框架。在实践中，已有一些考虑包容性的数据实践包括美国国立医学图书馆（National Library of Medicine，NLM）为方便所有的科研人员和社会公众都能无障碍地获取开放科学所公开的出版物、研究过程、数据和软件，积极采取措施，包括建设开放资源和数据中心、应用语义化技术、开发软件工具等，帮助人们共享开放科学成果。

（3）开放科学创新生态中的数据处理伦理更多强调“责任”。如何实现研究与创新的责任治理，引导科技创新朝着社会满意方向演进，是科技治理中的基本命题[47]。开放科学从诞生开始便蕴含主动开放的理念，包括科研人员、研究机构、政府、企业、公众等主体在参与开放科学的过程中都应具有伦理自觉性、主动承担相应责任，规范自身行为。在开放数据领域，开放科学框架（Open Science Framework）的在线社区，已形成数据信息开放的标准化管理与存储，可实现科研完整生命周期内数据资源的可读性与流通性，从而实现责任治理。在大数据、人工智能技术不断发展应用的背景下，更应强调数据处理的中的伦理责任，才能保障整个开放科学创新生态良好运转。2017年著名的《阿西洛马人工智能23条原则》、2018年出台的《欧盟人工智能》文件都体现了面对新兴科技伦理所体现出的责任意识，国内学者张梦在智能社会的基本伦理中明确提出“落实责任”这一内容。

（4）开放科学中的数据处理伦理准则也需要包含培训、监管和风险防控等内容。培训方面，数据处理伦理准则应面向参与开放科学中的所有参与主体制定培训计划，提升整体意识。监管方面，需要考虑在数据处理过程中，制定监管机构和执行机构，例如，墨尔本大学（The University of Melbourne，Australia）规定，科研中受试者的数据访问只有获得伦理委员会的许可才能执行，防止隐私泄露[48]。风险防控方面，应考虑如何从数据处理全流程出发，包括数据核验、记录、发布、权属、内容管理等多个环节，明确相应的责任和义务。

（5）在数据处理伦理中准则的实践落地方面，数据安全、准确、时效性、可信度，透明度等都开放科学中是必备条件。根据经济合作与发展组织（OECD）对开放科学的定义，开放科学强调主动向全社会披露科学知识、研究过程、研究数据与研究成果[10]，确保输出高质量、可信、可检验的信息，在此基础上才能实现知识共享与再利用，以及科学共同体的高度协同与合作。中国学者姚长青等认为，促进科研过程中的数据开放性和透明度有利于科研诚信建设[49]。

5 总 结

本研究面向开放科学创新生态的建设需求，引入数据管理领域中“数据处理伦理准则”这一重要概念，从共性与特性相结合的视角开展分析，总结归纳出数据处理伦理准则的一般维度，并结合开放科学理念和国家社会发展需要，提出包含3个维度15个要素的开

放科学创新生态的数据处理伦理准则框架，以期为更好理解把握开放科学创新生态中参与主体所应具备的伦理道德，以及如何制定开放科学创新生态中的数据管理实施方案提供了一定参考，未来可应用于制定适应中国开放科学创新生态发展的数据处理伦理准则文本等具体场景。但本文所提出的数据处理伦理准则框架，在准则内容、准则应用场景方面还有待结合开放科学创新生态的应用需求做进一步补充完善，这也正是本文不足之处，后期也将继续深化相关项研究。

参考文献：

[1] 黄金霞, 王昉, 姜恩波, 等. 融入开放科学生态的高端交流平台建设[J]. 数字图书馆论坛, 2021(12): 9-14.

HUANG J X, WANG F, JIANG E B, et al. Construction of high-end communication platform for open science ecology[J]. Digital library forum, 2021(12): 9-14.

[2] 中华人民共和国中央人民政府. 中共中央、国务院关于构建更加完善的要素市场化配置体制机制的意见 [EB/OL].(2020-04-09)[2022-08-25]. http://www.gov.cn/zhengce/2020-04/09/content_5500622.htm.

Government of the People's Republic of China. The central committee of the communist party of china and the state council's opinions on the construction of a more complete institutional mechanism for market-based allocation of factors[EB/OL]. (2020-04-09)[2022-08-25]. http://www.gov.cn/zhengce/2020-04/09/content_5500622.htm.

[3] 中华人民共和国政府. 国务院关于印发促进大数据发展行动纲要的通知[EB/OL]. (2015-09-05)[2022-08-25]. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm.

Government of the People's Republic of China. State Council's notice on the issuance of the outline of action to promote the development of big data[EB/OL]. (2015-09-05)[2022-08-25]. http://www.gov.cn/zhengce/content/2015-09/05/content_10137.htm.

[4] 黄金霞, 汪煊, 杨恒, 等. 来自多方的声音: 构建中国开放科学创新生态——第十届中国开放获取推介周(china OA week)会议综述[J]. 农业图书情报学报, 2022, 34(1): 49-61.

HUANG J X, WANG X, YANG H, et al. Summary of the 10th China OA week: Voices from some stakeholders on the construction of China's open scientific innovation ecology[J]. Journal of library and information science in agriculture, 2022, 34(1): 49-61.

[5] 习近平. 在中国科学院第二十次院士大会、中国工程院第十五次院士大会、中国科协第十次全国代表大会上的讲话[N]. 光明日报, 2021-05-29(2).

XI J P. Speech at the 20th congress of the Chinese academy of sciences, the 15th congress of the Chinese academy of engineering, and the 10th national congress of the Chinese association for science and technology[N]. Guangming ribao, 2021-05-29(2).

[6] DAMA 国际. DAMA 数据管理知识体系指南[M]. 北京: 机械工业出版社, 2020.

DAMA international. DMBOK2: Data management body of knowledge[M]. Beijing: China machine press, 2020.

[7] 陆康, 刘慧, 曹畋. 大数据技术视域下智慧图书馆伦理危机与控制研究[J]. 高校图书馆工作, 2021, 41(3): 46-50.

LU K, LIU H, CAO T. Research on ethical crisis and control of smart library from the perspective of big data technology[J]. Library work in colleges and universities, 2021, 41(3): 46-50.

[8] 张兴旺, 廖帅, 张鲜艳. 图书馆大数据资产的内涵、特征及其合理利用研究[J]. 情报理论与实践, 2019, 42(11): 15-20.

ZHANG X W, LIAO S, ZHANG X Y. Research on the connotation, characteristics and rational utilization of library's big data assets[J]. Information studies: Theory & application, 2019, 42(11): 15-20.

[9] 刘金亚, 顾立平, 张潇月, 等. 开放科研数据环境下科研人员的数据伦理框架研究[J]. 情报理论与实践, 2021, 44(2): 83-89.

LIU J Y, GU L P, ZHANG X Y. Framework of data ethics for researches in open research data environment[J]. Information studies: Theory & application, 2021, 44(2): 83-89.

[10] 陈劲, 阳镇. 数字化时代下的开放科学: 伦理难题与推进路径[J]. 吉林大学社会科学学报, 2021, 61(3): 116-128, 236.

CHEN J, YANG Z. Open science in digital era: Concept connotation, ethical difficulties and promotion path[J]. Jilin university journal social sciences edition, 2021, 61(3): 116-128, 236.

[11] 李伦. "楚门效应"数据巨机器的"意识形态"——数据主义与基于权利的数据伦理[J]. 探索与争鸣, 2018(5): 29-31.

LI L. The "ideology" of the "truman effect" data mega-machine: Dataism and rights-based data ethics[J]. Exploration and free views,

2018(5): 29-31.

[12] 链接人工智能准则平台[EB/OL]. [2022-09-20]. https://www.linking-ai-principles.org/.

Linking artificial intelligence principles[EB/OL].[2022-09-20]. https://www.linking-ai-principles.org/.

[13] 曾敬. 国外科研伦理研究及治理实践的思考——以“两用性”与“公众科学”问题为视角[J]. 中国高校科技, 2021(12): 45-49.

ZENG J. Reflections on foreign research ethics and governance practices: A perspective on "dual use" and "public science[J]. China university science & technology, 2021(12): 45-49.

[14] European union. General data protection regulation[EB/OL]. (2018-05-25)[2022-09-20]. https://gdpr-info.eu/.

[15] CANADA. Personal information protection and electronic documents act[EB/OL]. (2019-06-21)[2022-09-20]. https://laws-lois.justice.gc.ca/PDF/P-8.6.pdf.

[16] The United States departmeng of justics. Privacy act of 1974[EB/OL].(2021-04-30)[2022-09-20]. https://www.justice.gov/opcl/privacy-act-1974.

[17] S.2889-116th congress (2019-2020): A bill to safeguard data of americans from foreign governments that pose risks to national security by imposing data security requirements and strengthening review of foreign investments, and for other purposes[EB/OL]. (2019-11-18)[2022-09-20]. https://trackbill.com/bill/us-congress-senate-bill-2889-national-security-and-personal-data-protection-act-of-2019/1775410/.

[18] 中华人民共和国数据安全法 [EB/OL]. (2021-06-10)[2022-09-20]. http://www.npc.gov.cn/npc/c30834/202106/7c9af12f51334a73b56d7938f99a788a.shtml.

Data Security Law of the People's Republic of China [EB/OL]. (2021-06-10) [2022-09-20]. http://www.npc.gov.cn/npc/c30834/202106/7c9af12f51334a73b56d7938f99a788a.shtml.

[19] 中华人民共和国个人信息保护法 [EB/OL]. (2021-08-20)[2022-09-20]. http://www.npc.gov.cn/npc/c30834/202108/a8c4e3672c74491a80b53a172bb753fe.shtml.

Personal Information Protection Law of the People's Republic of China[EB/OL]. (2021-08-20)[2022-09-20]. http://www.npc.gov.cn/npc/c30834/202108/a8c4e3672c74491a80b53a172bb753fe.shtml.

[20] 中华人民共和国国家互联网信息办公室. 国家互联网信息办公室关于《数据出境安全评估办法(征求意见稿)》公开征求意见的通知 [EB/OL].(2021-10-29)[2022-09-20]. http://www.cac.gov.cn/2021-10/29/c_1637102874600858.htm.

Cyberspace Administration of China. Cyberspace administration of China's notice of public consultation on "data exit security assessment measures (draft for comments)"[EB/OL]. (2021-10-29)[2022-09-20]. http://www.cac.gov.cn/2021-10/29/c_1637102874600858.htm.

[21] 中华人民共和国国家互联网信息办公室. 国家互联网信息办公室关于《网络数据安全管理条例(征求意见稿)》公开征求意见的通知[EB/OL]. (2021-11-14)[2022-09-20]. http://www.cac.gov.cn/2021-11/14/c_1638501991577898.htm.

Cyberspace Administration of China. Cyberspace administration of China's notice of public consultation on "regulations on network data security management (draft for comments)" [EB/OL]. (2021-11-14)[2022-09-20]. http://www.cac.gov.cn/2021-11/14/c_1638501991577898.htm.

[22] The national commission for the protection of human subjects of biomedical and behavioral research. The belmont report[EB/OL]. (1979-04-18)[2022-09-20]. https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf.

[23] The department of homeland security. The menlo report: Ethical principles guiding information and communication technology research [EB/OL].(2012-08-03)[2022-09-20].https://www.dhs.gov/sites/default/files/publications/CSD-MenloPrinciplesCORE-20120803_1.pdf.

[24] The united nations development group. Data privacy, ethics and protection: Guidance note on big data for achievement of tht 2030 agenda[EB/OL].[2022-09-20]. https://unsdg.un.org/resources/data-privacy-ethics-and-protection-guidance-note-big-data-achievement-2030-agenda.

[25] Pernille tranberg, gry hasselbalch, catrine søndergaard byrne & birgitte kofod olsen. Dataethics - Principles and guidelines for companies, authorities & organisations[M]. Pernille publisher, 2020.

[26] European data protection supervisor. The EDPS ethics framework[EB/OL].(2019-11-12)[2022-09-20]. https://edps.europa.eu/data-pro-

tection/our-work/publications/other-documents/edps-ethics-framework_en.

[27] Central digital and data office. Data ethics framework [EB/OL]. (2018-06-13)[2022-09-20]. https://www.gov.uk/government/publications/data-ethics-framework.

[28] European commission. Ethics and data protection[EB/OL]. (2018-11-14)[2022-09-20]. https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-data-protection_en.pdf.

[29] The US general services administration. Federal data strategy data ethics framework [EB/OL].[2022-09-20]. https://strategy.data.gov/assets/docs/data-ethics-framework-action-14-draft-2020-sep-2.pdf.

[30] Swiss personalised health network. Ethical framework for responsible data processing in personalized health research[EB/OL]. (2018-05-07)[2022-09-20]. https://sphn.ch/wp-content/uploads/2019/11/Ethical_Framework_20180507_SPHN.pdf.

[31] 中华人民共和国科学技术部. 科学数据管理办法[EB/OL]. (2018-04-04)[2022-09-20]. https://most.gov.cn/xxgk/xinxifenlei/fdzdgknr/fgzc/gfxwj/gfxwj2018/201804/t20180404_139023.html.
Ministry of Science and Technology of the People's Republic of China. Scientific data management approach[EB/OL]. (2018-04-04)[2022-09-20]. https://most.gov.cn/xxgk/xinxifenlei/fdzdgknr/fgzc/gfxwj/gfxwj2018/201804/t20180404_139023.html.

[32] VICTORIA T, MARINA A, VASSO K, et al. Policy guidelines for open access and data dissemination and preservation [EB/OL]. (2013-02-01)[2022-09-20]. http://www.gsrt.gr/EOX/files/Policy%20guidelines%20for%20open%20access%20and%20data%20dissemination%20and%20preservation.pdf.

[33] Open data white paper unleashing the potential[EB/OL]. (2012-06-01) [2022-09-20].https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/78946/CM8353_acc.pdf.

[34] Open data policy - Managing information as an asset [EB/OL]. (2013-05-09)[2022-09-20].https://osec.doc.gov/opog/privacy/memorandums/omb_m-13-13.pdf.

[35] National institutes of health. Final NIH policy for data manager ent and sharing[EB/OL]. (2020-10-29)[2022-09-20]. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html.

[36] 马晓亭, 李强. 大数据环境下图书馆数据资源质量评估与整体优化研究[J]. 现代情报, 2017, 37(9): 103-106, 113.
MA X T, LI Q. Study for data source quality evaluation and global optimization for library in big data environment[J]. Journal of modern information, 2017, 37(9): 103-106, 113.

[37] Dataflux corporation. The data governance maturity model[EB/OL]. [2022-09-20]. https://www.fstech.co.uk/fst/whitepapers/The_Data_Governance_Maturity_Model.pdf.

[38] Gartner enterprise information & master data management summit. The gartner enterprise information management framework[EB/OL]. [2022-09-20].https://blogs.gartner.com/andrew_white/files/2016/10/On_site_poster.pdf.

[39] IBM. 数据治理 [EB/OL]. [2022-09-20]. https://www.ibm.com/cn-zh/analytics/data-governance.
IBM. Data governance [EB/OL].[2022-09-20]. https://www.ibm.com/cn-zh/analytics/data-governance.

[40] Stanford university. Data governance maturity model guiding questions for each Component-Dimension[EB/OL]. [2022-09-20]. http://www.teamnotes.fr/wp422/wp-content/uploads/2020/11/StanfordDataGovernanceMaturityModel.pdf.

[41] CMMI institute. Data management maturity (DMM)SM model[EB/OL]. [2022-09-20]. https://stage.cmmiinstitute.com/getattachment/cb35800b-720f-4afe-93bf-86ccefb1fb17/attachment.aspx.

[42] EDM council. DCAM v2 table of contents[EB/OL]. [2022-09-20]. https://cdn.ymaws.com/edmcouncil.org/resource/resmgr/DCAM_v2.2_Table_of_Contents_.pdf.

[43] 中国国家标准化管理委员会. 数据管理能力成熟度评估模型[EB/OL]. (2018-03-15)[2022-09-20]. http://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=B282A7BD34CAA6E2D742E0CAB7587DBC.
China national standardization administration committee. Data management capability maturity assessment model[EB/OL]. (2018-03-15)[2022-09-20]. http://openstd.samr.gov.cn/bzgk/gb/newGbInfo?hcno=B282A7BD34CAA6E2D742E0CAB7587DBC.

[44] UNESCO recommendation on open science [EB/OL].[2022-09-20]. https://en.unesco.org/science-sustainable-future/open-science/recommendation.

[45] 中华人民共和国科学技术部. 关于公开征求对《关于加强科技伦理治理的指导意见(征求意见稿)》意见的公告[EB/OL]. (2021-12-21)[2022-09-20]. https://www.most.gov.cn/tztg/202107/t20210728_176136.html.

Ministry of science and technology of the People's Republic of China. Announcement on public consultation on the guidance on strengthening ethical governance of science and technology (draft for comments) [EB/OL]. (2021-12-21) [2022-09-20]. https://www.most.gov.cn/tztg/202107/t20210728_176136.html.

[46] 杨巧云, 梁诗露, 杨丹. 数字包容: 发达国家的实践探索与经验借鉴[J]. 情报理论与实践, 2022, 45(3): 194-201.

YANG Q Y, LIANG S L, YANG D. Digital inclusion: Practical exploration and experience reference in developed countries[J]. Information studies: Theory & application, 2022, 45(3): 194-201.

[47] 梅亮, 臧树伟, 张娜娜. 新兴技术治理: 责任式创新视角的系统性评述[J]. 科学学研究, 2021, 39(12): 2113-2120, 2128.

MEI L, ZANG S W, ZHANG N N. The governance of emerging technology: A systematic review from responsible innovation perspective[J]. Studies in science of science, 2021, 39(12): 2113-2120, 2128.

[48] 叶冠成, 江雯欣, 代逸丹, 等. “开放科学”发展中的伦理问题探究——基于医学开放科学领域的分析[J]. 医学与哲学, 2019, 40(15): 32-36.

YE G C, JIANG W X, DAI Y D, et al. Exploring ethical problems in the development of "open science": Based on the analysis of open science in medicine[J]. Medicine & philosophy, 2019, 40(15): 32-36.

[49] 姚长青, 田瑞强. 开放科学中的数据诚信问题研究[J]. 科技与出版, 2019(1): 130-135.

YAO C Q, TIAN R Q. A study of data integrity issues in open science[J]. Science-technology & publication, 2019(1): 130-135.

Ethics Principle Framework of Data Handling for Open Scientific Innovation Ecology

LIU Jingyu[1,2], JIA Yujie[1,2*], HUANG Jinxia[1], WANG Fang[1]

(1. National Science Library, Chinese Academy of Sciences, Beijing 100190;

2. School of Economics and Management, University of Chinese Academy of Sciences, Beijing 100190)

Abstract: [Purpose/Significance] The development of open science still faces many data ethical problems, and there is an urgent need to explore the ethical guidelines for data handling that are compatible with the open science innovation ecology. The establishment of ethical guidelines is essential to create a favorable open science innovation ecology for data handling that match the construction of open science innovation ecology, and guide the participants in the open science innovation ecology to act in accordance with legal regulations and ethical norms in the process of data creation, sharing, dissemination, and utilization, and restraining data handling behaviors for the good and benefit of society. [Method/Process] This study first uses literature research and content analysis, based on relevant laws, guidelines, draft frameworks, policies and typical data management models on data handling ethics, from a universal perspective, based on the practice of data acquisition, storage, management, use and destruction, for social data in a broad sense, while taking into account a part of scientific research data, correlating and integrating different guidelines, seeking common themes and contents among existing guidelines, and summarizing the general dimensions and highly recognized guidelines. This study summarizes the general dimensions of

ethical guidelines for data handling and the contents of the guidelines with high recognition. Second, this study combines the concept of open science and the needs of national social development, and distills the needs of ethical guidelines for data handling in the open science innovation ecosystem from the latest domestic planning policies and the connotation of open science. Finally, this study synthesizes and compares the two results through matrix analysis, and analyzes the framework of ethical guidelines for data handling in the construction of open science innovation ecology from the perspectives of commonality and characteristics, and synergy between guideline contents and policies and concepts. [Results/Conclusions] The existing ethical guidelines for data handling have high applicability in the open science innovation ecology, but still need dynamic expansion. Combining the concept of open science and the development needs of national society, this study proposes a framework of ethical guidelines for data handling in the open science innovation ecology, including three dimensions and 15 elements of long-term governance, coordination and order, and systematic science, aiming to serve the further development of the text of ethical guidelines for data handling adapted to the development of the open science innovation ecology in China and to provide some reference for data management in institutions.

Keywords: data handling ethics; open and innovative ecology; open science; data handling; scientific ethics